

0590

1106

#4

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/880,729

DATE: 12/04/2001

TIME: 18:30:14

Input Set : N:\Crf3\RULE60\09880729.raw

Output Set: N:\CRF3\12042001\I880729.raw

SEQUENCE LISTING

ENTERED

3 (1) GENERAL INFORMATION:

4 (i) APPLICANT: Mathur, E., et al.

6 (ii) TITLE OF INVENTION: Carboxymethyl Cellulase from

7 Thermotoga Maritima

9 (iii) NUMBER OF SEQUENCES: 4

11 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

14 CECCHI, STEWART & OLSTEIN

15 (B) STREET: 6 BECKER FARM ROAD

16 (C) CITY: ROSELAND

17 (D) STATE: NEW JERSEY

18 (E) COUNTRY: USA

19 (F) ZIP: 07068

BEST AVAILABLE COPY

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: 3.5 INCH DISKETTE

23 (B) COMPUTER: IBM PS/2

24 (C) OPERATING SYSTEM: MS-DOS

25 (D) SOFTWARE: WORD PERFECT 5.1

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/880,729

C--> 29 (B) FILING DATE: 12-Jun-2001

30 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: US/09/472,857

34 (B) FILING DATE:

36 (A) APPLICATION NUMBER: 08/951,889

37 (B) FILING DATE:

39 (A) APPLICATION NUMBER: 08/518,615

40 (B) FILING DATE: August 23, 1995

42 (viii) ATTORNEY/AGENT INFORMATION:

43 (A) NAME: FERRARO, GREGORY D.

44 (B) REGISTRATION NUMBER: 36,134

45 (C) REFERENCE/DOCKET NUMBER: 331400-20

47 (ix) TELECOMMUNICATION INFORMATION:

48 (A) TELEPHONE: 201-994-1700

49 (B) TELEFAX: 201-994-1744

51 (2) INFORMATION FOR SEQ ID NO: 1:

53 (i) SEQUENCE CHARACTERISTICS:

54 (A) LENGTH: 954 BASE PAIRS

55 (B) TYPE: NUCLEIC ACID

56 (C) STRANDEDNESS: SINGLE

57 (D) TOPOLOGY: LINEAR

59 (ii) MOLECULE TYPE: cDNA

61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

63 ATGGGTGTTG ATCCTTTTGA AAGGAACAAA ATATTGGGAA GAGGCATTAA TATAGGAAAT 60

64 GCGCTTGAAG CACCAAATGA GGGAGACTGG GGAGTGGTGA TAAAAGATGA GTTCTTCGAC 120

RAW SEQUENCE LISTING

DATE: 12/04/2001

PATENT APPLICATION: US/09/880,729

TIME: 18:30:14

Input Set : N:\Crf3\RULE60\09880729.raw

Output Set: N:\CRF3\12042001\I880729.raw

```

65 ATTATAAAAG AAGCCGGTTT CTCTCATGTT CGAATTCCAA TAAGATGGAG TACGCACGCT 180
66 TACGCGTTTC CTCCTTATAA AATCATGGAT CGCTTCTTCA AAAGAGTGGA TGAAGTGATA 240
67 AACGGAGCCC TGAAGAGAGG ACTGGCTGTT GCTATAAATA TTCATCACTA CGAGGAGTTA 300
68 ATGAATGATC CAGAAGAACA CAAGGAAAGA TTTCTTGCTC TTTGGAAACA AATTGCTGAT 360
69 CGTTATAAAG ACTATCCCGA AACTCTATTT TTTGAAATTC TGAATGAACC TCACGGAAAT 420
70 CTTACTCCGG AAAAATGGAA TGAAGTCTT GAGGAAGCTC TAAAAGTTAT AAGATCAATT 480
71 GACAAAAGC ACACTATAAT TATAGGCACA GCTGAATGGG GGGGTATATC TGCCCTTGAA 540
72 AAAGTGTCTG TCCCAAAATG GGAAAAAAT TCTATAGTTA CAATTCACCTA CTACAATCCT 600
73 TTCGAATTTA CCCATCAAGG AGCTGAGTGG GTGGAAGGAT CTGAGAAATG GTTGGGAAGA 660
74 AAGTGGGGAT CTCCAGATGA TCAGAAACAT TTGATAGAAG AATTCAATTT TATAGAAGAA 720
75 TGGTCAAAAA AGAACAAAAG ACCAATTTAC ATAGGTGAGT TTGGTGCCTA CAGAAAAGCT 780
76 GACCTTGAAT CAAGAATAAA ATGGACCTCC TTTGTCGTTT GCGAAATGGA GAAAAGGAGA 840
77 TGGAGCTGGG CATACTGGGA ATTTTGTTC GGTTTTGGTG TTTATGATAC TCTGAGAAAA 900
78 ACCTGGAATA AAGATCTTTT AGAAGCTTTA ATAGGAGGAG ATAGCATTGA ATAA 954

```

80 (2) INFORMATION FOR SEQ ID NO: 2:

81 (i) SEQUENCE CHARACTERISTICS:

82 (A) LENGTH: 317 AMINO ACIDS

83 (B) TYPE: AMINO ACID

84 (C) STRANDEDNESS:

85 (D) TOPOLOGY: LINEAR

87 (ii) MOLECULE TYPE: PROTEIN

89 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

91 Met Gly Val Asp Pro Phe Glu Arg Asn Lys Ile Leu Gly Arg Gly Ile
92      5              10              15
93 Asn Ile Gly Asn Ala Leu Glu Ala Pro Asn Glu Gly Asp Trp Gly Val
94      20              25              30
95 Val Ile Lys Asp Glu Phe Phe Asp Ile Ile Lys Glu Ala Gly Phe Ser
96      35              40              45
97 His Val Arg Ile Pro Ile Arg Trp Ser Thr His Ala Tyr Ala Phe Pro
98      50              55              60
99 Pro Tyr Lys Ile Met Asp Arg Phe Phe Lys Arg Val Asp Glu Val Ile
100 65              70              75              80
101 Asn Gly Ala Leu Lys Arg Gly Leu Ala Val Ala Ile Asn Ile His His
102      85              90              95
103 Tyr Glu Glu Leu Met Asn Asp Pro Glu Glu His Lys Glu Arg Phe Leu
104      100             105             110
105 Ala Leu Trp Lys Gln Ile Ala Asp Arg Tyr Lys Asp Tyr Pro Glu Thr
106      115             120             125
107 Leu Phe Phe Glu Ile Leu Asn Glu Pro His Gly Asn Leu Thr Pro Glu
108      130             135             140
109 Lys Trp Asn Glu Leu Leu Glu Glu Ala Leu Lys Val Ile Arg Ser Ile
110 145             150             155             160
111 Asp Lys Lys His Thr Ile Ile Ile Gly Thr Ala Glu Trp Gly Gly Ile
112      165             170             175
113 Ser Ala Leu Glu Lys Leu Ser Val Pro Lys Trp Glu Lys Asn Ser Ile
114      180             185             190
115 Val Thr Ile His Tyr Tyr Asn Pro Phe Glu Phe Thr His Gln Gly Ala
116      195             200             205
117 Glu Trp Val Glu Gly Ser Glu Lys Trp Leu Gly Arg Lys Trp Gly Ser

```

BEST AVAILABLE COPY

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/880,729

DATE: 12/04/2001

TIME: 18:30:14

Input Set : N:\Crf3\RULE60\09880729.raw

Output Set: N:\CRF3\12042001\I880729.raw

```

118      210      215      220
119 Pro Asp Asp Gln Lys His Leu Ile Glu Glu Phe Asn Phe Ile Glu Glu
120 225      230      235      240
121 Trp Ser Lys Lys Asn Lys Arg Pro Ile Tyr Ile Gly Glu Phe Gly Ala
122      245      250      255
123 Tyr Arg Lys Ala Asp Leu Glu Ser Arg Ile Lys Trp Thr Ser Phe Val
124      260      265      270
125 Val Arg Glu Met Glu Lys Arg Arg Trp Ser Trp Ala Tyr Trp Glu Phe
126      275      280      285
127 Cys Ser Gly Phe Gly Val Tyr Asp Thr Leu Arg Lys Thr Trp Asn Lys
128      290      295      300
129 Asp Leu Leu Glu Ala Leu Ile Gly Gly Asp Ser Ile Glu
130 305      310      315

```

132 (2) INFORMATION FOR SEQ ID NO: 3:

134 (i) SEQUENCE CHARACTERISTICS:

135 (A) LENGTH: 51 BASE PAIRS

136 (B) TYPE: NUCLEIC ACID

137 (C) STRANDEDNESS: SINGLE

138 (D) TOPOLOGY: LINEAR

W--> 140 (ii) MOLECULE TYPE: Oligonucleotide

142 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

144 TTATTGCGGC CGCTTAAGGA GGAAAAAATT ATGGGTGTTG ATCCTTTTGA A 51

146 (2) INFORMATION FOR SEQ ID NO: 4:

148 (i) SEQUENCE CHARACTERISTICS:

149 (A) LENGTH: 33 BASE PAIRS

150 (B) TYPE: NUCLEIC ACID

151 (C) STRANDEDNESS: SINGLE

152 (D) TOPOLOGY: LINEAR

W--> 154 (ii) MOLECULE TYPE: Oligonucleotide

156 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

157 TTATTGGATC CGAAGGTTGA AACCACGCCA TCT

33

BEST AVAILABLE COPY

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/880,729

DATE: 12/04/2001

TIME: 18:30:15

Input Set : N:\Crf3\RULE60\09880729.raw

Output Set: N:\CRF3\12042001\I880729.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:140 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3

L:154 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4

BEST AVAILABLE COPY